DOCUMENT RESUME

ED 451 934 PS 029 395

AUTHOR Shaul, Marnie S.

TITLE Early Childhood Programs: The Use of Impact Evaluations To Assess Program Effects. Report to the Chairman, Subcommittee on Oversight of Government Management, Restructuring and the District of Columbia, Committee on Governmental Affairs, U.S. Senate.
INSTITUTION General Accounting Office, Washington, DC.
REPORT NO GAO-01-542
PUB DATE 2001-04-00
NOTE 27p.
AVAILABLE FROM U.S. General Accounting Office, P.O. Box 37050, Washington, DC 20013. Tel: 202-512-6000; Fax: 202-512-6061; e-mail: Info@www.gao.gov; Web site: http://www.gao.gov.
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Day Care; Evaluation Problems; *Outcomes of Education; *Preschool Education; Program Effectiveness; Program Evaluation; *Research Design; Research Problems; Scientific Methodology; Special Education
IDENTIFIERS Child Care and Development Block Grants; Even Start; Migrant Education Program; Project Head Start; Temporary Assistance for Needy Families

ABSTRACT

At the request of a Senate subcommittee, this report describes the value of conducting impact evaluations, describes their current use in evaluating selected early childhood education and care programs, and discusses the value of other types of early childhood education and care studies currently promoted and sponsored by the Departments of Health and Human Services (HHS) and Education. Current government-sponsored studies were reviewed for 11 programs that fund early childhood education and care and that serve at least 25,000 children under age 5, including Project Head Start, Even Start, Migrant Education, and State, Preschool, and Infant and Toddler Grants for special education. The report notes that impact evaluations are considered by many to be the best method for determining the extent to which the program, rather than other factors, causes participant outcomes. The design of impact evaluation studies is discussed, with focus on methodological rigor, expense and time required, and ethical issues involved. Findings noted indicate that the two federal programs focused most on early childhood education, Head Start and Even Start, are currently being studied using impact evaluations with experimental designs. One HHS study of Head Start is a 6-year national impact study that will follow children through first grade. Evaluation of the Early Head Start program will collect information about 17 local programs and follow children to their third birthday. The 6-year Even Start evaluation will examine the effectiveness of 18 local programs. The report notes that although Congress mandated study of Even Start's effectiveness, it did not require that an impact evaluation be used. For the other nine programs, impact evaluations are not currently being used to study the effectiveness of the early childhood education and care services they support. The report concludes by noting that HHS and Education promote and sponsor many types of research and evaluation studies, thereby

Reproductions supplied by EDRS are the best that can be made from the original document.

providing agencies with answers to a broad range of questions about program operation. (KB)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 451 934

PS

United States General Accounting Office

GAO

Report to the Chairman, Subcommittee on Oversight of Government Management, Restructuring and the District of Columbia, Committee on Governmental Affairs, U.S. Senate

April 2001

EARLY CHILDHOOD PROGRAMS

The Use of Impact Evaluations to Assess Program Effects

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PS 029395

GAO
Accountability * Integrity * Reliability

BEST COPY AVAILABLE

GAO-01-542

BEST COPY AVAILABLE

Contents

Letter 1

Appendix I **GAO Contacts and Staff Acknowledgments** 21

Table

Table 1: Characteristics of Programs Supporting Early Childhood Education and Care, Fiscal Year 1999 5

Figures

Figure 1: Children Are Exposed to Many Factors That Affect Their Development 9
Figure 2: How Experimental Designs Work 10
Figure 3: Categorization of the Early Childhood Education and Care Programs Included in Our Review 17

Abbreviations

CCDF	Child Care and Development Fund
FACES	Family and Child Experiences Survey
GPRA	Government Performance and Results Act
HHS	Department of Health and Human Services
SSBG	Social Services Block Grant
TANF	Temporary Assistance for Needy Families

GAO
Accountability * Integrity * Reliability

United States General Accounting Office
Washington, DC 20548

April 16, 2001

The Honorable George V. Voinovich
Chairman, Subcommittee on Oversight of
Government Management, Restructuring
and the District of Columbia
Committee on Governmental Affairs
United States Senate

Dear Mr. Chairman:

In 1999, at least $9 billion in federal funds supported early education and care services for children younger than 5.[1] These funds provided a variety of services and support to children and their families to meet a range of goals, from increasing low-income families' access to affordable child care to ensuring that children's educational, health, and social needs are met. Given the large federal investment and the current attention on the importance of early childhood experiences, policymakers are interested in the effectiveness of federally funded early childhood education and care efforts and the types of research designs available for determining whether these programs are meeting their objectives for children.

Because of your particular interest in the use of impact evaluations as a way of determining program effectiveness, our objectives were to (1) describe the value of conducting impact evaluations, (2) describe their current use in evaluating selected early childhood education and care programs, and (3) discuss the value of other types of early childhood education and care studies the Departments of Health and Human Services (HHS) and Education currently promote and sponsor.

To understand the use of impact evaluations, we reviewed current government-sponsored studies (both agency-initiated and congressionally mandated) of 11 programs that fund early childhood education and care and report serving 25,000 or more children younger than age 5. Thus, we examined current evaluations and evaluation proposals for two programs focused on early childhood education—Head Start and Even Start—and for the following nine programs: Child Care and Development Fund

[1] *Early Education and Care: Overlap Indicates Need to Assess Crosscutting Programs* (GAO/HEHS-00-78, Apr. 28, 2000).

(CCDF), Temporary Assistance for Needy Families (TANF), Twenty-First Century Community Learning Centers, three special education programs (State, Preschool, and Infant and Toddler Grants), Social Services Block Grant (SSBG), Migrant Education, and Title I.[2] These nine programs provide early childhood education and care though, for most, it is not a primary activity. We conducted our work between October 2000 and March 2001 in accordance with generally accepted government auditing standards.

Results in Brief

Many researchers consider impact evaluations to be the best method for determining the extent to which the program itself, rather than other factors, is causing participant outcomes. Impact evaluations can be designed in several ways, but fall into two basic design categories: experimental and quasiexperimental. Experimental designs randomly assign eligible individuals either to a group that will receive services from the program being studied or to a group that will not receive services from the program. The relevant outcomes of these two groups are measured and compared, and any differences found between the two can be attributed to the programs. While rigorous, these evaluations can raise ethical issues because they deny services to eligible individuals. This issue, along with other issues characteristic of complex evaluations, can make experimental impact evaluations especially difficult, time-consuming, and expensive to perform. Quasiexperimental designs use a comparison group that is not created with random assignment. As a result, the comparison group's similarities to program participants are less certain. Quasiexperimental designs are less rigorous, but tend to be less expensive and time-consuming and do not have the same ethical issues.

The two federal programs focused most on early childhood education, Head Start and Even Start, are currently being studied using impact evaluations with experimental designs. Both of these programs are intended to produce improvements in children's school readiness and educational outcomes, including enhanced literacy. HHS is conducting two Head Start studies, both of which were mandated by the Congress. The first is a 6-year national impact study of 3-to-4-year-olds that will follow children through first grade. This evaluation will cost about $28.3 million. The second is of the Early Head Start program. This $21 million study will collect information about 17 local programs and follow children

[2]Title I refers only to title I, part A, of the Elementary and Secondary Education Act.

up to their third birthday. HHS has initiated another study that will follow these children through the spring before their entry into kindergarten. The $3.6-million Even Start study will examine the effectiveness of 18 local programs and will take 6 years to complete. The Congress mandated a study of Even Start's effectiveness but did not require that effectiveness be assessed using an impact evaluation. For the other nine programs we reviewed, impact evaluations are not currently being used to study the effectiveness of the early childhood education and care services they support.[3] Researchers have used a variety of study designs, other than impact evaluations, to understand more about these types of programs with respect to early childhood education and care.

HHS and Education promote and sponsor many types of research and evaluation studies. The value of a varied study agenda is that it provides agencies with answers to a broad range of questions about program operation and allows them to align research with the focus of the program. For example, HHS and Education sponsor evaluations that provide information on whether programs are being implemented as intended. The Early Head Start study currently under way includes an implementation study that provides information on the program's development and services provided to children and their families. For its CCDF program—which is aimed at increasing the availability, affordability, and quality of child care services—HHS believes that increasing the level of knowledge and information about how child care systems work is a critical step toward improving the availability and quality of child care services. Therefore, HHS sponsors basic research to better understand the complex nature of child care markets at the state, local, and national levels.

Background

The use of impact evaluations to assess the effect of a program or other intervention is an accepted practice in areas as diverse as medicine, economics, and social services. Some frequently cited examples of impact evaluations involve experiments designed to test the efficacy of drugs by randomly assigning participants to a group that will either receive or not receive the experimental drug. Impact evaluations are also used in other areas to assess program effectiveness. One such study is currently under way for the Job Corps. In addition, several state-level impact evaluations

[3]For TANF, several impact studies are under way that will examine the effect of TANF work requirements on children's outcomes; however, the effectiveness of early childhood education and care services is not the primary focus of these evaluations.

are being used to study welfare-to-work programs. The Congress has mandated studies to ensure that agencies conduct the kinds of evaluations it deems necessary. In both authorizing and reauthorizing legislation, the Congress has required evaluations of a program or demonstration program.

Impact evaluations have been used to evaluate and develop programs in the early childhood education and care arena. For example, HHS sponsored an impact evaluation of a demonstration project called the Comprehensive Child Development program to test the effectiveness of the program for low-income families and their young children. HHS incorporated many of the lessons learned from this study in the Early Head Start program.

The 11 federal programs that serve 25,000 or more young children in some capacity are administered by either Education or HHS (table 1). The programs differ in a number of ways, including their goals and the portion of funds they devote to young children. For example, Head Start's goal is school readiness and most of its budget is spent on children younger than age 5. All Head Start grantees provide services that are designed to enhance school readiness. In contrast, SSBG, which devotes only 11 percent of its funds for early childhood education and care, is a flexible source of funds that allows states wide discretion to fund a variety of social services, one of which can be child care. Its goal is primarily to help families achieve self-sufficiency. The other programs we reviewed fund early childhood education and care in various ways but often have primary goals that are only indirectly related to children or school readiness.

Table 1: Characteristics of Programs Supporting Early Childhood Education and Care, Fiscal Year 1999

Program	Agency	Program goals	Target age	Major program emphasis for children[a]	Proportion of children under 5 receiving early childhood education or child care services		Percentage of budget authority allocated to children under 5	Estimated number of participants under age 5[b]
					Education	Child care		
Head Start[c]	HHS	Promote school readiness by enhancing the social and cognitive development of low-income preschool children through the provision of health, educational, nutritional, social, and other services that are determined to be necessary	Birth through 5[d]	Early childhood education	All or nearly all	Some	94	780,200
Even Start—State Educational Agencies	Education	Improve the educational opportunities of low-income families by integrating early childhood education, adult literacy or adult basic education, and parenting education into a unified family literacy program	Birth through 7	Early childhood education	All or nearly all	Some	.. [e]	25,500
Special Education—Grants for Infants and Toddlers With Disabilities	Education	Ensure that early intervention services are provided to infants and toddlers with disabilities that will minimize their potential for developmental delay	Birth through 2	Early intervention services	Most	None	100	186,819
Special Education—Preschool Grants	Education	Ensure that 3- through 5-year-old children with disabilities have access to a free and appropriate public education	3 through 5	Special Education and related services	All or nearly all	[f]	55	316,000

Program	Agency	Program goals	Target age	Major program emphasis for children[a]	Proportion of children under 5 receiving early childhood education or child care services		Percentage of budget authority allocated to children under 5	Estimated number of participants under age 5[b]
					Education	Child care		
Special Education—Grants to States	Education	Ensure that 3- through 21-year-old individuals with disabilities have access to a free and appropriate public education	3 through 21	Special Education and related services	All or nearly all	[c]	6	574,713
Title I	Education	Improve the teaching and learning of children who are failing, or are most at risk of failing, to meet challenging state academic standards	Birth to 21	Elementary and secondary education	All or nearly all	Few	5[d]	313,000[e]
Migrant Education—Basic State Grant Program	Education	Supplement existing education efforts to ensure that migrant children have the opportunity to meet the same state educational standards as other children	Birth through 21	Elementary and secondary education	Some	Few	13	69,251
CCDF	HHS	Increase the availability, affordability, and quality of child care services	Birth through 12	Child care	Some	All or nearly all	70[f]	1,260,000
Twenty-First Century Community Learning Centers	Education	Enable public elementary and secondary schools to plan, implement, or expand extended learning opportunities for the benefit of the educational, health, social service, cultural, and recreational needs of their communities	Not specified	After-school learning opportunities	Some	Few	[g]	1,760,000

Program	Agency	Program goals	Target age	Major program emphasis for children[a]	Proportion of children under 5 receiving early childhood education or child care services		Percentage of budget authority allocated to children under 5	Estimated number of participants under age 5[b]
					Education	Child care		
SSBG	HHS	Achieve or maintain economic self-support and self-sufficiency; prevent or remedy neglect, abuse, and exploitation of children and adults; and prevent or reduce inappropriate institutional care	Not specified	Child care and other social services	Not known	Not known	11	2,320,067[f]
TANF	HHS	Provide assistance for needy families; end the dependence of needy parents by promoting job preparation, work, and marriage; prevent and reduce out-of-wedlock pregnancies; and encourage the formation and maintenance of two-parent families	Not specified	Child care and other purposes	Not known	Most	7.5	350,000[f]

[a]Although some of these programs provide a range of services for children and their families, for this table we focused only on the programs' emphasis on young children.

[b]We did not compute a total because children may have participated in more than one of the programs listed. For example, a 3-year-old with a disability may have received services under Education's Special Education–Preschool Grants as well as under its Special Education–Grants to States program. Moreover, services for children may be supported with CCDF and SSBG funds as well.

[c]The Early Head Start program is a component of Head Start. Early Head Start specifically provides services for pregnant women and families with infants and toddlers.

[d]Head Start is authorized to serve children at any age before the age of compulsory school attendance; however, most children are below age 5.

[e]Agency officials were unable to estimate the amount spent on children under age 5.

[f]The program was not allowed to provide this service.

[g]This figure is an estimate based on school year 1999–2000 data from a GAO survey.

BEST COPY AVAILABLE

[1]This percentage is an estimate based on the number of children under 5 receiving subsidies.

[2]Agency officials do not collect participation data for children under age 5. The reported estimate includes all children receiving child care, a high proportion of whom are assumed to be under age 5.

[3]This number is based on monthly TANF enrollment data and includes the estimated number of children under 5 years old receiving child care through TANF as well as the number of children in this age group who receive child care paid for with TANF funds transferred to CCDF.

The passage of the Government Performance and Results Act (GPRA) of 1993 influenced how agencies assess the effectiveness of their programs. GPRA shifted the focus of accountability for federal programs from inputs, such as staffing and activity levels, to outcomes. GPRA requires that each federal agency develop a multiyear strategic plan identifying the agency's mission and long-term goals. To develop the plans, each agency must connect its long-term strategic goals with daily activities. While GPRA requires agencies to measure the outcomes of their programs, it does not require agencies to conduct formal program evaluations such as impact evaluations. However, GPRA requires agencies to summarize the findings of program evaluations in their annual performance reports.

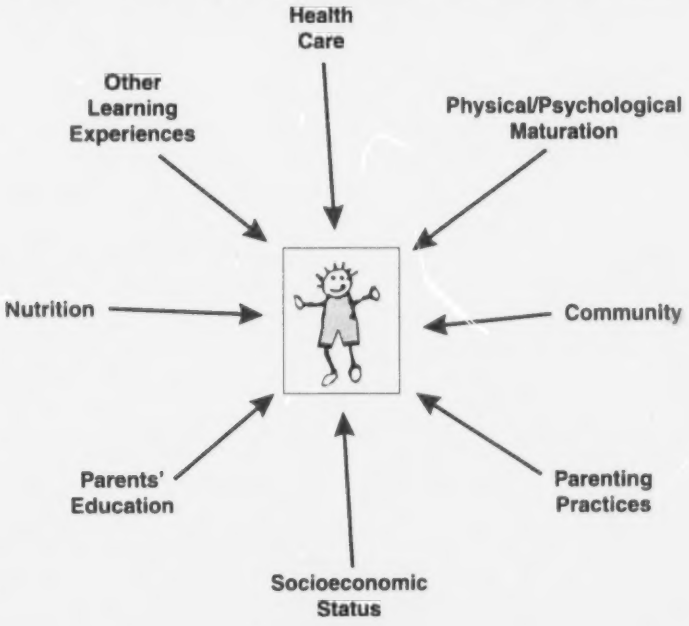
Impact Evaluations Are the Most Definitive Method for Determining a Program's Effect on Its Participants

Many researchers consider impact evaluations to be the best method for determining a program's effect on its participants because they isolate a program's contribution from the effects of other factors that could have influenced participant outcomes. Impact evaluations can be designed in several ways, but most fall into two categories: experimental and quasiexperimental designs.

Impact evaluations are used to isolate the influences of a program being studied from other influences in an individual's life.[4] In the case of a child, many influences affect his or her development (fig. 1). As a child grows, he or she acquires new knowledge and new skills. Nutrition, health, family, and community as well as education and care play roles in his or her learning. In light of all these influences, it becomes difficult to distinguish between the effects of the program and the other factors that influence a child's learning.

[4]For a discussion of impact evaluations, see Carol H. Weiss, *Evaluation Research: Methods of Assessing Program Effectiveness* (Englewood Cliffs, N.J.: Prentice-Hall Inc., 1972); Robert F. Boruch, *Randomized Experiments for Planning and Evaluation: A Practical Guide* (Thousand Oaks, Calif.: Sage Publications, 1997).

Figure 1: Children Are Exposed to Many Factors That Affect Their Development



To isolate the program's influences on a child, an impact evaluation studies two groups of children: those receiving program services and a similar group not receiving program services. Researchers compare the relevant outcomes, such as reading ability, of these two groups of children to determine the program's effect.

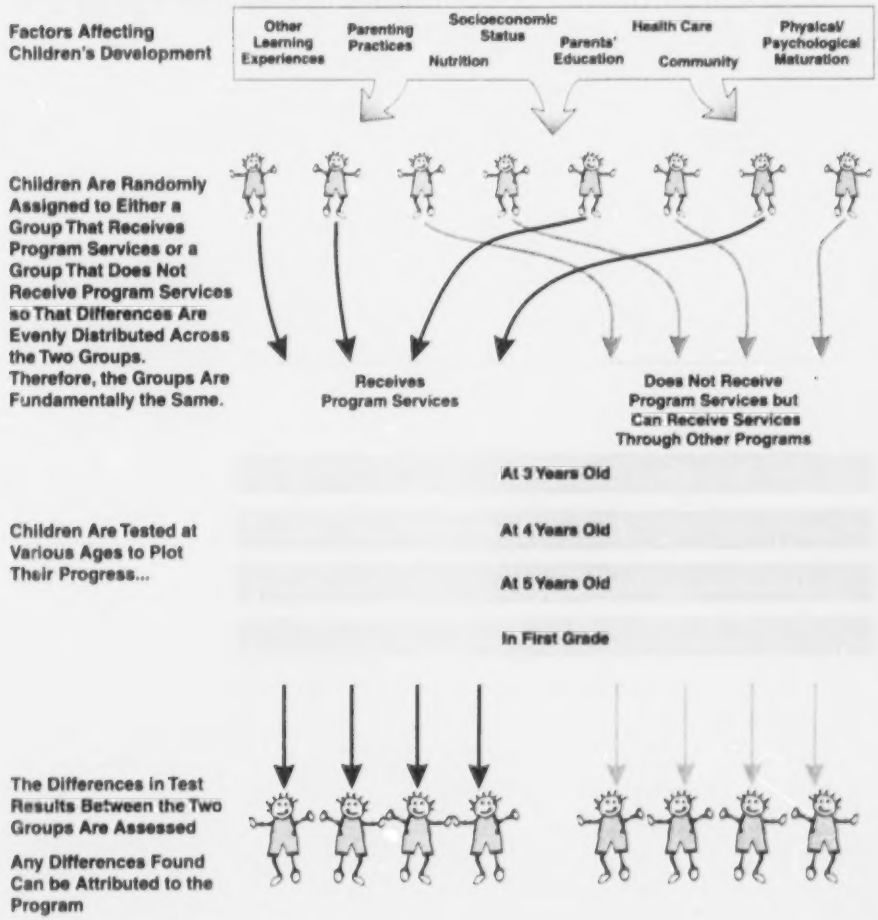
Most impact evaluations are one of two types—experimental or quasiexperimental—although they can be designed in several ways. The two designs differ primarily in the way that the comparison groups are developed. In an experimental design, the comparison group is referred to as the "control" group. It is composed of individuals randomly selected from possible program participants. In a quasiexperimental impact evaluation, the comparison group is composed of individuals who share characteristics with program participants, but who may or may not have ever sought program services.

Experimental Design

The more rigorous of the two types of impact evaluations is the experimental design (fig. 2). It randomly assigns individuals to either a group that will receive program services or a group that will not receive

program services. As applicants apply for program services, each has an equal chance of being placed in a group that will receive program services or a group that will not receive program services. Although individual applicants are different, random assignment distributes those differences equally between the two groups, resulting in two groups that are presumed to have no systematic differences in their characteristics. Although individuals in the control group do not receive services from the program under study, they may receive similar services through other programs.

Figure 2: How Experimental Designs Work


Factors Affecting Children's Development
Other Learning Experiences
Parenting Practices
Socioeconomic Status
Nutrition
Parents' Education
Health Care
Community
Physical/ Psychological Maturation
Children Are Randomly Assigned to Either a Group That Receives Program Services or a Group That Does Not Receive Program Services so That Differences Are Evenly Distributed Across the Two Groups. Therefore, the Groups Are Fundamentally the Same.
Receives Program Services
Does Not Receive Program Services but Can Receive Services Through Other Programs
Children Are Tested at Various Ages to Plot Their Progress...
At 3 Years Old
At 4 Years Old
At 5 Years Old
In First Grade
The Differences in Test Results Between the Two Groups Are Assessed
Any Differences Found Can be Attributed to the Program

At various points during an experimental impact study, researchers measure the relevant outcomes of the members of both groups. The group that is not receiving services demonstrates what would have happened without the program. Thus, when the progress of the two groups is compared, any differences between them can be attributed to the influences of the program. This way, the use of a control group serves to rule out alternative explanations for differences in the outcomes of program participants and nonparticipants. However, because control group members may receive services similar to those provided through the program under study, the differences in outcomes between the two groups may be less than they would have been had the control group not received those similar services.

The Denial of Services to Certain Eligible Study Participants Heightens the Complexity and Difficulty of Experimental Impact Evaluations

The denial of the program services under study to eligible individuals in control groups—a necessary aspect of experimental impact evaluations—raises a number of issues. For example, experimental designs cannot be used when a program must serve all eligible individuals. The denial of services can also raise ethical issues, the result of which can be reluctance on the part of program staff and officials to participate in impact studies. These issues, along with other challenges characteristic of complex evaluations, can make experimental impact evaluations especially difficult, time-consuming, and expensive to perform.

Ethical issues arise in the selection of an experimental impact evaluation control group. As we stated earlier, a control group consists of participants who are denied services. Several factors may complicate the creation of the control group, and in fact denying services for the purpose of creating a control group may not be possible. For example, a program may be required to serve all eligible individuals, as is the case with Special Education programs. Even when forming a control group does not raise these kinds of issues, program staff and officials may be ethically opposed to denying services to individuals, especially for evaluations requiring several years to complete. To overcome such opposition, an additional investment of time may be necessary to gain the trust of program staff. The organization conducting the evaluation may need to provide training to staff so that they understand the need for random assignment and can collect data if necessary.

In addition to the possible opposition of staff members, people seeking program services can be reluctant to agree to forego program participation. Researchers must gain the acceptance and informed consent of both participants and control group members (or, in the case of evaluations involving minors, their guardians). In some instances, control

group members or their families may require monetary or other incentives to participate. Over the course of the evaluation, incentives may need to be sustained or increased.

The ethical issues associated with experimental designs, along with other challenges characteristic of complex evaluations, can make them difficult to complete and add to the time and expense they require. Experimental impact evaluations are often complex, multiyear studies awarded through competitive bids to experienced research firms. Considerable attention must be given to both study planning and execution. For example, researchers often prepare detailed plans describing the study, its methodologies, and issues that can affect the study. In many cases, aspects of the study must be pilot tested before the larger study can be undertaken. Then, once evaluations are under way, individuals may need to be tracked over several years. For example, an evaluation that begins well before a child starts school—for example, at age 3 years—may require 3 years or more of data collection to track the child through the beginning of his or her school years. In addition, special efforts must be made to limit the number of study participants who leave the study, especially from within the control group.

Quasiexperimental Design

The other, less rigorous type of impact evaluation is quasiexperimental design. When randomly assigning individuals to a control group is not a feasible option, quasiexperimental impact evaluations can be used to compare the performance of program participants on various measures to individuals not in the program. In a quasiexperimental design, methods other than random assignment are used to create a comparison group. A comparison group can be developed in a variety of ways. One way is to use a set of individuals who have similar characteristics to the group receiving the program services under study. For example, the group might live in the same neighborhood as the group receiving services and have family and income characteristics similar to those receiving serivces.

The extent to which quasiexperimental comparison groups are actually similar to program participants is not entirely known. For this reason, quasiexperimental evaluations cannot rule out all of the factors that influence participants' outcomes. Other factors could explain differences between the two groups, making it difficult to conclude, with certainty, the effect of the program being examined. When well planned and executed, however, such designs can provide some indication of program impact.

Impact Evalutions Are Being Used to Assess the Effectiveness of Early Education Programs

The two programs most focused on early childhood education—Head Start and Even Start—are currently being studied using impact evaluations with an experimental design. HHS is conducting impact studies of Head Start and Early Head Start, and Education is using an impact evaluation to study Even Start. Both of these programs are intended to improve children's school readiness and educational and developmental outcomes, for example by encouraging enhanced literacy. Impact evaluations are not currently being used to evaluate the early education and care services provided by the other nine programs we reviewed whose goals are not as directly focused on early childhood education as Even Start and Head Start. For many of these programs, providing early childhood education and care is one of many allowable services, not the primary program focus.

The Head Start Study

The Head Start evaluation is a $28.3-million, nationally representative, longitudinal impact evaluation. The contract for the evaluation was awarded in fall 2000 to Westat Inc., a research firm, in collaboration with the Urban Institute, the American Institute for Research, and Decision Information Resources. The evaluation is still in its early stages and data collection has not yet commenced. Before data collection begins in fall 2002, a pilot study will be conducted to test various procedures and measures and to investigate the variations that exist across different communities with respect to Head Start programs and the availability of other care options for low-income children. The pilot study will be conducted in spring 2001.

In fall 2002, researchers will begin to collect data on about 5,000 to 6,000 3- and 4-year-olds from 75 programs and communities across the country; data collection will continue through the spring of the children's first grade of elementary school. It is estimated that this evaluation will take more than 6 years to complete. The Congress mandated this evaluation and specified that it be an impact evaluation by stating that children should be randomly assigned to either a group that will receive Head Start services or one that will not receive Head Start services.

The study has two primary goals. The first is to determine the national impact of Head Start on children's school readiness by comparing children in Head Start to children not in Head Start. Data in areas related to school readiness, such as cognition, physical well-being, social and emotional development, and language usage, will be collected from both groups of children. In addition, information on parenting practices, family structure, demographics, and socioeconomics will be collected. The second goal is

to determine under which conditions and for which children Head Start works best.

In crafting the study proposal, Westat noted the difficulties inherent in denying services to eligible children in order to create the control group and suggested ways to soften the ethical objections to random assignment. For example, the proposal allows Head Start staff to provide information to control group families on alternative services in the community, and exempts a small number of children with significant disabilities from the random assignment. The plan also acknowledges the need to make parents comfortable with the process and, in some instances, provides financial incentives for participation. At the same time, Westat noted the need to maintain the integrity of the random assignment over a number of years. To this end, the plan calls for providing information to Head Start staff about the need for random assignment. In addition, Head Start grantees will need to agree to exclude children assigned to the control group from enrolling at other Head Start centers within their jurisdictional control during the study. The final report is due in December 2006.

The issue of Head Start's effectiveness was highlighted in 1997 when we reported that the body of research on the Head Start program was insufficient to draw conclusions about its national impact.[5] The report recommended that a study or studies be undertaken to determine the impact of the program nationally. The following year, we reiterated these points in another report and added that the federal government's significant financial investment in Head Start, including plans to increase the number of children served, warranted definitive evaluation studies, even though they may be costly.[6]

Based on these recommendations and the testimony of research methodologists and early childhood experts, the Congress mandated that HHS fund and conduct an evaluation to determine, on a national level, the impact of Head Start on the children it serves. The Congress also called for an expert panel to develop recommendations regarding the study design to determine if, overall, Head Start programs' impacts are consistent with the primary goal of school readiness. The panel included 30 experts in the

[5] *Head Start: Research Provides Little Information on Impact of Current Program* (GAO/HEHS-97-59, Apr. 15, 1997).

[6] *Head Start: Challenges in Monitoring Program Quality and Demonstrating Results* (GAO/HEHS-98-186, June 30, 1998).

areas of program evaluation and research, education, early childhood care and care policy, and economics. The panel outlined a framework for the impact study and deliberated extensively on the feasibility, ethics, and credibility of random assignment. It concluded that random assignment represented the best approach for answering the question of Head Start's impact.

The Early Head Start Study

The Early Head Start program is part of Head Start and serves pregnant women, infants, and toddlers. The Early Head Start Study, which was also congressionally mandated, began in 1995 shortly after the program was established. It will cost about $21 million[7] and take 6 years to complete. This longitudinal study of about 3,000 families and their children will assess children 14, 24, and 36 months after birth. The Early Head Start study includes an impact component as well as an implementation study that we discuss later in this report. Mathematica Policy Research, Inc. (a research organization), and researchers at Columbia University are conducting the study in collaboration with 15 local research teams.

The impact portion of the study will examine Early Head Start programs at 17 local sites that include a mix of urban and rural sites as well as sites serving individuals with different racial and ethnic backgrounds. These 17 sites were selected to reflect the characteristics of all programs funded in the early years of the program.

As the Early Head Start grantees recruited families, the families were randomly assigned to either a group that will receive Early Head Start program services or a control group that will not receive program services. (Control group families could receive other services in the community.) The final report, which will follow children up to their third birthday, is due in 2002. HHS has initiated another study which will follow these children through the spring before their entry into kindergarten to evaluate the program's impact on school readiness, among other things. This follow-up evaluation was not congressionally mandated.

The Even Start Study

The estimated $3.6 million Even Start study will also examine several local projects using an experimental design. The study is supported through

[7]This estimate includes funding for the impact evaluation and the implementation study that are part of a larger Early Head Start study.

program funds that the Congress made available for evaluation. The mandate for the Even Start study was included in the program's original legislation and, although that legislation has been amended several times, the requirement for an independent evaluation to determine the performance and effectiveness of Even Start has remained unchanged.

Funded in 1997, the study provides for an experimental design to test the effectiveness of 18 projects serving 400 Even Start families and 200 control group families. The contract, awarded to Abt Associates, Inc. (a research organization), will, among other things, measure children's school readiness. The study will take 6 years to complete and the final report is expected in summer 2003.

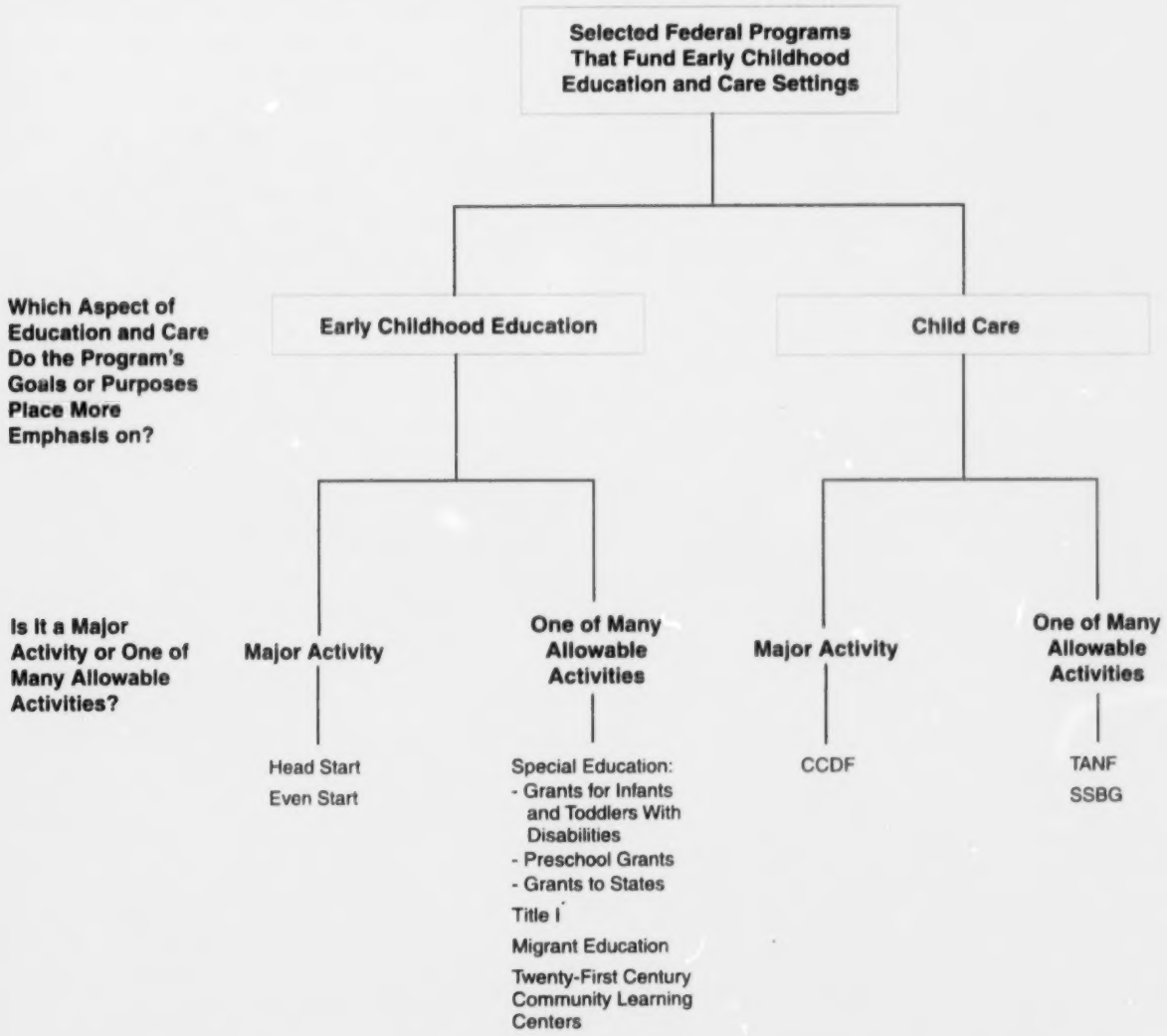
The current study is actually the second Even Start impact study conducted using an experimental design. The first evaluation included an impact evaluation that examined Even Start programs operated by five grantees.[8] The small number of sites examined by the study and the lack of information on control group experiences did not permit conclusions about program effectiveness.

Impact evaluations are not currently being used to evaluate the early education and care services provided by the other nine programs we reviewed.[9] Their goals are not as focused on early childhood education as those of Even Start and Head Start. For many of these programs, providing early childhood education and care is one of many allowable services, rather than the primary program focus (fig. 3). Researchers have used a variety of study designs, other than impact evaluations, to understand more about these types of programs with respect to early childhood education and care services.

[8]Abt Associates, Inc., and RMC Research Corporation, *National Evaluation of the Even Start Family Literacy Program* (Washington, D.C.: Jan. 23, 1995).

[9]For TANF, several impact studies are under way that will examine the effect of TANF work requirements on children's outcomes; however, the effectiveness of early childhood education and care services is not the primary focus of these evaluations.

Figure 3: Categorization of the Early Childhood Education and Care Programs Included in Our Review


Selected Federal Programs That Fund Early Childhood Education and Care Settings
Which Aspect of Education and Care Do the Program's Goals or Purposes Place More Emphasis on?
Early Childhood Education
Child Care
Is It a Major Activity or One of Many Allowable Activities?
Major Activity
One of Many Allowable Activities
Major Activity
One of Many Allowable Activities
Head Start
Even Start
Special Education:
- Grants for Infants and Toddlers With Disabilities
- Preschool Grants
- Grants to States
Title I
Migrant Education
Twenty-First Century Community Learning Centers
CCDF
TANF
SSBG

Note: Figure adapted from GAO report entitled, *Early Childhood Programs: Characteristics Affect the Availability of School Readiness Information* (GAO/HEHS-00-38, Feb. 28, 2000). This figure illustrates the major emphasis of the programs, although some programs provide both child care and early childhood education, both of which can contribute to positive outcomes for children. For the purposes of this illustration, we categorized the programs according to their main emphasis.

Different Types of Studies Inform HHS and Education About Administered Programs

HHS and Education promote and sponsor many types of studies, other than impact evaluations, that can provide a wide variety of data valuable to program managers and policymakers. Different study designs are used depending on the questions to be answered, the nature of the program being studied, and the type of information needed. Often, to answer varied, complex, and interrelated questions, policymakers may need to use several different designs to assess a single program. Policymakers' desire to know a program's effect does not preclude their need to know about its implementation and other details that can be revealed through different types of studies.[10]

Implementation Evaluation

An implementation evaluation, also called a process evaluation, can assess the extent to which a program is being administered as the Congress anticipated. An implementation evaluation typically assesses whether program activities conform to statutory and regulatory requirements, program design, and professional standards or customer expectations. Generally, an implementation evaluation necessitates collecting information for comparison to legislative objectives. For example, if legislation requires a program to serve a particular population with certain characteristics, researchers conducting an implementation evaluation would examine the demographics of participants to determine if the program is indeed serving this population.

Implementation evaluations provide information on many aspects of a program. For example, the first Even Start study was largely an implementation study, and its findings were used to improve program design and shape future research. The study was broad in scope, but examined the characteristics of Even Start participants, projects, and services to assess how closely they resembled what had been envisioned for the program. The study served as a catalyst for changes in the program's legislation, including a shift in focus on those most in need. In addition, as a result of the study's findings, teen parents are now eligible for program services.

The Early Head Start study under way also includes an implementation component. The implementation study gathered information on

[10]See *Performance Measures and Evaluation: Definitions and Relationships* (GAO/GGD-98-26, Apr. 1998) for a discussion of research designs.

participants and services to determine if the program was implemented as intended. Information from the implementation study will be integrated with the results of the impact analysis.

Outcome Evaluation

An outcome evaluation assesses how closely a program's achievements are aligned with program goals. This type of evaluation answers the question, "Are participants achieving intended outcomes?"

Outcome evaluations can determine how well children are doing at a particular time. However, without the use of an experimental design, researchers cannot attribute this performance to any one influence. Unlike an impact evaluation, outcome evaluations are unable to isolate the influences of the program under study.

HHS and Education use outcome evaluations to find out how well children in agency-sponsored programs perform basic skills. For example, Education is sponsoring two national studies with outcome components for its special education programs. One study is an early intervention study. This study will follow a nationally representative sample of children from birth to age 3, monitoring them during and after the program, to describe their progress in such areas as eating and dressing. The other study follows 3- to 5-year-olds through early elementary school and will also address questions about children's outcomes. HHS funded the Family and Child Experiences Survey (FACES) in 1996 to collect outcome data from a nationally representative sample of Head Start programs, children, and families. FACES collects a range of data that includes cognitive, social, emotional, and physical development of Head Start children; the well-being and accomplishments of Head Start families; and the quality of Head Start classrooms.

Research Studies

Research studies are used to collect information to improve an agency's understanding about a particular area, such as early childhood education, or the issues surrounding a particular area.

Often, research studies are used to test ideas that have arisen in an area but are not yet proven. For example, HHS believes that increasing the knowledge of how child care systems work is a critical step toward improving the availability and quality of child care services. To explore whether this is the case, HHS is using research studies to understand the complex nature of child care markets at the state, local, and national levels. Thus, some of the questions framing HHS' child care research agenda include: What does child care look like today? How do child care

variations influence child development and well-being? How do child care variations affect family economic self-sufficiency and well-being?

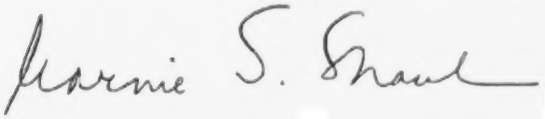
Agency Comments

We provided a draft of this report to the Departments of Education and Health and Human Services for comment, and they did not provide comments. However, Education's Office of Elementary and Secondary Education and Office of Special Education and Rehabilitative Services and HHS' Administration for Children and Families provided technical comments, which we incorporated where appropriate.

We are sending copies of this report to the Honorable Richard Durbin, Ranking Member of the Subcommittee on Oversight of Government Management, Restructuring and the District of Columbia, Committee on Governmental Affairs; the Honorable Roderick R. Paige, Secretary of Education; the Honorable Tommy G. Thompson, Secretary of Health and Human Services; relevant congressional committees, and other interested parties. We will also make copies available to others on request.

Please contact me on (202) 512-7215 if you or your staff have any questions about this report. Other GAO contacts and staff acknowledgments are listed in appendix I.

Sincerely yours,



Marnie S. Shaul
Director, Education, Workforce
and Income Security Issues

Appendix I: GAO Contacts and Staff Acknowledgments

GAO Contacts

Karen Whiten, (202) 512-7291
Sherri Doughty, (202) 512-7273

Staff Acknowledgments

In addition to those named above, John Smale and Corinna Nicolaou made key contributions to this report.

Ordering Information

The first copy of each GAO report is free. Additional copies of reports are $2 each. A check or money order should be made out to the Superintendent of Documents. VISA and MasterCard credit cards are accepted, also.

Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:
U.S. General Accounting Office
P.O. Box 37050
Washington, DC 20013

Orders by visiting:
Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders by phone:
(202) 512-6000
fax: (202) 512-6061
TDD (202) 512-2537

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

Orders by Internet:
For information on how to access GAO reports on the Internet, send an e-mail message with "info" in the body to:

info@www.gao.gov

or visit GAO's World Wide Web home page at:

http://www.gao.gov

To Report Fraud, Waste, or Abuse in Federal Programs

Contact one:

- **Web site: http://www.gao.gov/fraudnet/fraudnet.htm**
- **e-mail: fraudnet@gao.gov**
- **1-800-424-5454 (automated answering system)**

United States
General Accounting Office
Washington, D.C. 20548-0001

Presorted Standard
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested

BEST COPY AVAILABLE